Journal of Number Theory 129 (2009) 580–587

Contents lists available at ScienceDirect

Journal of Number Theory

www.elsevier.com/locate/jnt

The fluctuations in the number of points on a hyperelliptic curve over a finite field

Pär Kurlberg [a,*,1], Zeév Rudnick [b,2]

[a] Department of Mathematics, Royal Institute of Technology, SE-100 44 Stockholm, Sweden
[b] Raymond and Beverly Sackler School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel

ARTICLE INFO

Article history:
Received 4 May 2008
Revised 26 August 2008
Available online 21 October 2008
Communicated by J. Brian Conrey

ABSTRACT

The number of points on a hyperelliptic curve over a field of q elements may be expressed as $q+1+S$ where S is a certain character sum. We study fluctuations of S as the curve varies over a large family of hyperelliptic curves of genus g. For fixed genus and growing q, Katz and Sarnak showed that $S/\sqrt{q}$ is distributed as the trace of a random $2g\times 2g$ unitary symplectic matrix. When the finite field is fixed and the genus grows, we find that the limiting distribution of S is that of a sum of q independent trinomial random variables taking the values ± 1 with probabilities $1/2(1+q^{-1})$ and the value 0 with probability $1/(q+1)$. When both the genus and the finite field grow, we find that $S/\sqrt{q}$ has a standard Gaussian distribution.

1. Introduction

Given a finite field $\mathbb{F}_q$ of odd cardinality q and a square-free monic polynomial $F\in\mathbb{F}_q[X]$ of degree $d\geqslant 3$, we get a smooth projective hyperelliptic curve C_F with affine model

$$C_F\colon\ Y^2=F(X)$$

having genus $g=(d-2)/2$ when d is even and $g=(d-1)/2$ when d is odd. In this note we study the fluctuations in the number of $\mathbb{F}_q$-points on C_F when F is drawn at random from the set of all

* Corresponding author.
E-mail addresses: kurlberg@math.kth.se (P. Kurlberg), rudnick@post.tau.ac.il (Z. Rudnick).

[1] The author was partially supported by grants from the Göran Gustafsson Foundation, the Knut and Alice Wallenberg foundation, the Royal Swedish Academy of Sciences, and the Swedish Research Council.
[2] The author was supported by the Israel Science Foundation (grant No. 925/06).

0022-314X/$ – see front matter
doi:10.1016/j.jnt.2008.09.004

square-free monic polynomials $F \in \mathbb{F}_q[X]$ of degree d, where the probability measure is obtained by picking the coefficients of F uniformly in $\mathbb{F}_q^d$ and conditioning on F being square free. Correspondingly we get a probability measure on a family of hyperelliptic curves of genus $g \geqslant 1$ defined over $\mathbb{F}_q$. Our goal is to study these fluctuations in the limit of either large genus or large q, or both.

The number of $\mathbb{F}_q$-points on C_F can be written as[3] $q + S(F) + 1$ where $S(F)$ is the character sum

$$S(F) = \sum_{x \in \mathbb{F}_q} \chi\big(F(x)\big)$$

and χ is the quadratic character of $\mathbb{F}_q^\times$ (with the convention that $\chi(0) = 0$). Thus the problem is equivalent to studying the fluctuations of $S(F)$ as F varies over all square-free polynomials in $\mathbb{F}_q[X]$ of degree d, in the limit as either d or q (or both) grow.

Our finding is that there are three distinct types of distribution results according to the way the parameters g and q are allowed to grow:

(i) For q fixed and the genus $g \to \infty$, we find that $S(F)$ is distributed asymptotically as a sum of q independent identically distributed (i.i.d.) trinomial random variables $\{X_i\}_{i=1}^q$, i.e., random variables taking values in 0, ± 1 with probabilities $1/(q+1)$, $1/2(1+q^{-1})$ and $1/2(1+q^{-1})$, respectively.
(ii) When the genus g is fixed and $q \to \infty$, $S(F)/\sqrt{q}$ is distributed as the trace of a random matrix in the group $USp(2g)$ of $2g \times 2g$ unitary symplectic matrices. This is due to Katz and Sarnak [5].
(iii) When both $g \to \infty$ and $q \to \infty$ we find that $S(F)/\sqrt{q}$ has a Gaussian value distribution with mean zero and variance unity.

The case (iii) when both variables grow can be thought of as a limiting case of either the two previous ones, when one of the two parameters is held fixed. It is thus a good consistency check to see that the limit distributions in both cases (i) and (ii) are a standard Gaussian. Indeed, in the case when q is fixed, (i) gives that the limit distribution of $S(F)/\sqrt{q}$ is that of a normalized sum $(X_1 + \cdots + X_q)/\sqrt{q}$ of q i.i.d. random variables; in turn the distribution of a normalized sum of such i.i.d.'s converges, as $q \to \infty$, to a Gaussian distribution with mean zero and variance unity by the Central Limit Theorem. In case the genus g is fixed, (ii) gives that the limit distribution of $S(F)/\sqrt{q}$ is that of the traces of random matrices in $USp(2g)$. The limit distribution of traces of a random matrix in $USp(2g)$, as $g \to \infty$, is a standard Gaussian by a theorem of Diaconis and Shahshahani [3]. Of course this is not a proof of (iii), as it only addresses the limiting form of the *limit distribution* in (i) and (ii), that is either $\lim_{q\to\infty}(\lim_{g\to\infty})$ or $\lim_{g\to\infty}(\lim_{q\to\infty})$ and not the joint limit $\lim_{q\to\infty, g\to\infty}$.

1.1. Some related work

1. In the unpublished manuscript [8], Larsen studied moments for a related family of hyperelliptic curves, namely curves of the form $Y^2 = \prod_{i=1}^n (X - a_i)$, where $a_1, \ldots, a_n$ ranges over all n-tuples consisting of distinct elements of $\mathbb{F}_q$, and obtained Gaussian moments.
2. Knizhnerman and Sokolinskii [6,7] computed moments of the character sum $S(F)$ when F ranges over all monic non-square (rather than square-free) polynomials to investigate extreme values taken by $S(F)$ (we thank Igor Shparlinski for this reference).
3. Bergström [1] used methods closely related to ours in order to obtain equivariant point counts for families of hyperelliptic curves. These point counts were then used to determine cohomology groups of the moduli space of stable curves of genus 2 with n marked points, for $n \leqslant 7$.

[3] Giving the number of points of C_F a spectral interpretation, it is more natural to write the number of points as $q - S'(F) + 1$, where $S'(F) = -S(F)$ is the trace of the action induced by the Frobenius automorphism on a certain cohomology group. However, for our purposes, studying $S(F)$ will lead to slightly simpler notation, and, as we shall see, the distribution of $S(F)$ is symmetric, hence $S(F)$ and $S'(F)$ have the same distribution.

4. Finally, we refer to the recent preprint of Faifman and Rudnick [4] which studies the statistics of the zeros of the zeta function of the curves C_F over a fixed finite field in the limit of large genus.

1.2. The main results

Before giving a more quantitative statement of our main results, we will need some notation. Let $V_d \subset \mathbb{F}_q[X]$ be the set of monic polynomials of degree d, and let $\mathcal{F}_d \subset V_d$ be the subset of *square-free* polynomials of degree d. We will model $S(F)$ as a sum of q independent identically distributed (i.i.d.) trinomial random variables $\{X_i\}_{i=1}^q$, where each X_i takes values in $0, \pm 1$ with probabilities $1/(q+1)$, $1/2(1+q^{-1})$ and $1/2(1+q^{-1})$, respectively.

For q fixed and $d \to \infty$, we show that $S(F)$ behaves as $\sum_{i=1}^q X_i$ in the following sense:

Theorem 1. *If q is fixed and d tends to infinity then the distribution of $S(F)$, as F ranges over all elements in $\mathcal{F}_d$, is that of a sum of q independent trinomial random variables. More precisely, for $s \in \mathbb{Z}$ with $|s| \leqslant q$, we have*[4]

$$\frac{|\{F \in \mathcal{F}_d\colon\ S(F) = s\}|}{|\mathcal{F}_d|} = \text{Prob.}\left(\sum_{i=1}^{q} X_i = s\right) \cdot \left(1 + O\left(q^{(3q-d)/2}\right)\right).$$

Remark. We may also let q tend to infinity in Theorem 1, provided that d tends to infinity in such a way that $d > 3q$.

By studying the moments we find that $S(F)/\sqrt{q}$ has a Gaussian value distribution when both d, q tend to infinity.

Theorem 2. *If d, q both tend to infinity, then the moments of $S(F)/\sqrt{q}$ are asymptotically Gaussian with mean* 0 *and variance* 1. *In particular the limiting value distribution is a standard Gaussian.*

2. Proof of Theorem 1

The idea of the proof is to make the following heuristic precise: Putting the uniform probability measure on $\mathcal{F}$, we may view $f \to S(f)$ as a random variable on $\mathcal{F}$. $S(f)$ can in turn be written as

$$S(f) = \sum_{x \in \mathbb{F}_q} X_x,$$

where for each $x \in \mathbb{F}_q$, $X_x = \chi(f(x))$ is also a random variable on $\mathcal{F}$. Then, as d grows, the variables $\{X_x\}_{x \in \mathbb{F}_q}$ become independent and the distribution of each individual X_x is that of the earlier mentioned trinomial random variable.

Thus, we will study the following slightly more general problem: Given a subset $S \subset \mathbb{F}_q$ and a tuple $a = (a_x)_{x \in S}$, $a_x \in \mathbb{F}_q$, we wish to find the probability that for a randomly selected $F \in \mathcal{F}$ we have $F(x) = a_x$ for all $x \in S$.

Before proceeding we need to introduce some additional notation. For $F \in \mathbb{F}_q[X]$, write $F = \prod_{i=1}^n F_i^{e_i}$ as a product of irreducible polynomials, and let

$$\mu(F) := \begin{cases} 0 & \text{if } e_i > 1 \text{ for some } i, \\ (-1)^n & \text{if } F \text{ is square free.} \end{cases}$$

Further, put

[4] Here, and in what follows, all constants implied by the $O(\cdot)$-notation will be absolute.

$$|F| := q^{\deg(F)}$$

and let

$$\zeta_q(s) := \sum_{F \text{ monic}} |F|^{-s} = \prod_{F \text{ irreducible}} \left(1 - |F|^{-s}\right)^{-1} = \frac{1}{1 - q^{1-s}}$$

be the (incomplete) zeta function of $\mathbb{A}^1/\mathbb{F}_q$.

We will need to know the number of square-free monic polynomials, which can easily be deduced from the identity

$$\zeta_q(s) = \zeta_q(2s) \sum_{d \geqslant 0} |\mathcal{F}_d| q^{-ds}, \qquad \Re(s) > 1.$$

Lemma 3. *The number of square-free monic polynomials of degree d equals*

$$|\mathcal{F}_d| = \begin{cases} q^d - q^{d-1} = q^d/\zeta(2) & \text{if } d \geqslant 2, \\ q^d & \text{if } d = 0, 1. \end{cases}$$

We shall also need the following simple counting lemma which is at the heart of the independence result.

Lemma 4. *For $l \leqslant q$ let $x_1, x_2, \ldots, x_l \in \mathbb{F}_q$ be distinct elements, and let $a_1, a_2, \ldots, a_l \in \mathbb{F}_q$. If $d \geqslant l$, then*

$$\left|\left\{F \in V_d\colon\ F(x_1) = a_1, \ldots, F(x_l) = a_l\right\}\right| = q^{d-l}.$$

Proof. Let $\tilde{V}_d = \{g \in \mathbb{F}_q[X]\colon\ \deg(g) \leqslant d-1\}$. The map $f(X) \to g(X) := f(X) - X^d$ then defines a bijection from V_d to $\tilde{V}_d$. Since $f(x_i) = a_i$ for $1 \leqslant i \leqslant l$, is equivalent to $g(x_i) = a_i - x_i^d$ for $1 \leqslant i \leqslant l$, we find that

$$\left|\left\{f \in V_d\colon\ f(x_i) = a_i \text{ for } 1 \leqslant i \leqslant l\right\}\right| = \left|\left\{g \in \tilde{V}_d\colon\ g(x_i) = a_i - x_i^d \text{ for } 1 \leqslant i \leqslant l\right\}\right|. \tag{2.1}$$

Now, the evaluation map $g \to (g(x_1), \ldots, g(x_l))$ is a linear map from $\tilde{V}_d$ to $\mathbb{F}_q^l$, and its kernel consists of all $g \in \tilde{V}_d$ that are divisible by $\prod_{i=1}^{l}(x - x_i)$. Hence the $\mathbb{F}_q$-dimension of the kernel equals $d - l$, and since $\dim_{\mathbb{F}_q}(\tilde{V}_d) = d$ the cokernel has dimension l. In particular, the evaluation map is surjective, and both sides of (2.1) equal q^{d-l} for all choices of $a_1, \ldots, a_l$. □

Next we determine the probability of a random polynomial in $\mathcal{F}_d$ taking a prescribed set of *nonzero* values on l points.

Lemma 5. *Let $d \geqslant 2$ and $l \leqslant q$ be a positive integers, let $x_1, x_2, \ldots, x_l \in \mathbb{F}_q$ be distinct elements, and let $a_1, a_2, \ldots, a_l \in \mathbb{F}_q$ be* nonzero *elements. Then*

$$\frac{|\{F \in \mathcal{F}_d\colon\ F(x_1) = a_1, F(x_2) = a_2, \ldots, F(x_l) = a_l\}|}{|\mathcal{F}_d|} = \frac{q^{-l}}{(1 - q^{-2})^l} \cdot \left(1 + O\left(q^{l-d/2}\right)\right).$$

Proof. Using inclusion–exclusion, we find that

$$\begin{aligned}&\left|\left\{F\in\mathcal{F}_d\colon\ F(x_i)=a_i \text{ for } 1\leqslant i\leqslant l\right\}\right|\\&\quad=\sum_{\substack{F\in V_d: F(x_i)=a_i\\ \text{for } 1\leqslant i\leqslant l}}\mu(F)^2=\sum_{D:\deg(D)\leqslant d/2}\mu(D)\left|\left\{F\in V_{d-2\deg(D)}\colon\ D(x_i)^2F(x_i)=a_i \text{ for } 1\leqslant i\leqslant l\right\}\right|.\end{aligned}$$

With $\sum'$ denoting the sum over all polynomials such that $D(x)\neq 0$ for all $x\in\{x_1,x_2,\dots,x_l\}$, we find, since $a_i\neq 0$ for all $i\leqslant l$, that the above equals

$$\sum_{D:\deg(D)\leqslant d/2}{}'\mu(D)\left|\left\{F\in V_{d-2\deg(D)}\colon\ F(x_i)=a_iD(x_i)^{-2} \text{ for } 1\leqslant i\leqslant l\right\}\right|. \tag{2.2}$$

Now, as long as $\deg(F)=d-2\deg(D)\geqslant l$, by Lemma 4, we have

$$\left|\left\{F\in V_{d-2\deg(D)}\colon\ F(x_i)=a_iD(x_i)^{-2} \text{ for } 1\leqslant i\leqslant l\right\}\right|=q^{d-2\deg(D)-l}$$

hence (2.2) equals

$$q^{d-l}\sum_{D:\deg(D)<(d-l)/2}{}'\mu(D)q^{-2\deg(D)}+\text{Error} \tag{2.3}$$

where, since there can be at most one polynomial F of degree smaller than l that attains l prescribed values (at l distinct points),

$$\text{Error}\ll\sum_{D:(d-l)/2\leqslant\deg(D)\leqslant d/2}1=O\left(q^{d/2}\right).$$

Our next goal is to evaluate the main term

$$\sum_{D:2\deg(D)<d-l}{}'\mu(D)q^{-2\deg(D)}=\sum_{D}{}'\mu(D)q^{-2\deg(D)}+O\left(q^{(l-d)/2}\right).$$

Noting that

$$\begin{aligned}\sum_{D}{}'\mu(D)|D|^{-2s}&=\prod_{\substack{F:F \text{ is irreducible},\\ F(x_i)\neq 0 \text{ for } i\leqslant l}}\left(1-|F|^{-2s}\right)\\&=\left(1-q^{-2s}\right)^{q-l}\cdot\prod_{\substack{F:F \text{ is irreducible}\\ \deg(F)>1}}\left(1-|F|^{-2s}\right)=\frac{1}{\zeta(2s)(1-q^{-2s})^l}\end{aligned}$$

we find that (2.3) equals

$$q^{d-l}\left(\frac{1}{\zeta(2)(1-q^{-2})^l}+O\left(q^{(l-d)/2}\right)\right)+O\left(q^{d/2}\right)=\frac{q^{d-l}}{\zeta(2)(1-q^{-2})^l}+O\left(q^{d/2}\right). \tag{2.4}$$

Since $|\mathcal{F}_d|=\frac{q^d}{\zeta(2)}$ for $d\geqslant 2$, the probability that $F(x_i)=a_i$ for all $i\leqslant l$ equals

$$\frac{q^{-l}}{(1-q^{-2})^l} + O\big(q^{-d/2}\big),$$

concluding the proof. □

We now easily obtain the probability of F attaining any set of prescribed values.

Proposition 6. *Let $x_1, x_2, \dots, x_l, x_{l+1}, x_{l+m} \in \mathbb{F}_q$ be distinct elements, let $a_1, a_2, \dots, a_l \in \mathbb{F}_q^\times$, and let $a_{l+1} = a_{l+2} = \cdots = a_{l+m} = 0$. Then*

$$\frac{1}{|\mathcal{F}_d|}\big|\{F \in \mathcal{F}_d\colon\ F(x_i) = a_i \text{ for } 1 \leqslant i \leqslant m+l\}\big| = \frac{(1-1/q)^m q^{-(m+l)}}{(1-q^{-2})^{m+l}} \cdot \big(1 + O\big(q^{(3m+2l-d)/2}\big)\big). \quad (2.5)$$

Proof. Any $F \in \mathcal{F}_d$ which vanishes at $Z = \{x_{l+1}, x_{l+2}, \dots, x_{l+m}\}$ can be written as

$$F(x) = \prod_{i=l+1}^{l+m} (x - x_i) G(x)$$

where $G \in \mathcal{F}_{d-m}$ is a square-free polynomial that is non-vanishing on Z. Moreover, the condition that $F(x_i) = a_i$ for $1 \leqslant i \leqslant l$, can then be expressed as $G(x_i) = a_i \prod_{j=l+1}^{l+m} (x_i - x_j)^{-1}$ for $1 \leqslant i \leqslant l$, and $G(x_j)$ is arbitrary (but nonzero) for $l+1 \leqslant j \leqslant l+m$. In other words, there are $(q-1)^m$ possible values for G restricted to Z and by Lemma 5 (in particular, see (2.4)), the number of such polynomials equals

$$(q-1)^m\left(\frac{q^{d-m-(m+l)}}{\zeta(2)(1-q^{-2})^{m+l}} + O\big(q^{(d-m)/2}\big)\right) = (1-1/q)^m\left(\frac{q^{d-(l+m)}}{\zeta(2)(1-q^{-2})^{m+l}} + O\big(q^{(d+m)/2}\big)\right).$$

Dividing by the number of square-free polynomials, we find that the probability of a random $F \in \mathcal{F}_d$ vanishing on Z, and taking prescribed values outside Z equals

$$\frac{(1-1/q)^m q^{-(m+l)}}{(1-q^{-2})^{m+l}}\big(1 + O\big(q^{(3m+2l-d)/2}\big)\big). \qquad \square$$

To finish the proof of Theorem 1, we argue as follows: Let $x_1, x_2, \dots, x_q$ be distinct elements of $\mathbb{F}_q$, let $\epsilon_i \in \{-1, 0, 1\}$ for $1 \leqslant i \leqslant q$, and define $m = |\{i \in \{1, 2, \dots, q\}\colon\ \epsilon_i = 0\}|$. Taking $l = q - m$ in Proposition 6 and noting that the number of nonzero squares, respectively non-squares, in $\mathbb{F}_q$ equals $(q-1)/2$, we find that

$$\begin{aligned}
&\frac{|\{F \in \mathcal{F}_d\colon\ \chi(F(x_i)) = \epsilon_i \text{ for all } 1 \leqslant i \leqslant q\}|}{|\mathcal{F}_d|} \\
&= \left(\frac{q-1}{2}\right)^{q-m} \cdot \frac{(1-1/q)^m q^{-q}}{(1-q^{-2})^q}\big(1 + O\big(q^{(3q-d)/2}\big)\big) \\
&= 2^{-(q-m)}\frac{(1-1/q)^q q^{-m}}{(1-q^{-2})^q}\big(1 + O\big(q^{(3q-d)/2}\big)\big) = \frac{2^{-(q-m)} q^{-m}}{(1+q^{-1})^q}\big(1 + O\big(q^{(3q-d)/2}\big)\big).
\end{aligned}$$

On the other hand, if $X_1, \dots, X_q$ are i.i.d. trinomial random variables as before, we have

$$\text{Prob.}(X_i = \epsilon_i \text{ for } 1 \leqslant i \leqslant q) = (q+1)^{-m} \cdot 2^{-(q-m)}\big(1 + q^{-1}\big)^{m-q} = 2^{-(q-m)}\big(1 + q^{-1}\big)^{-q} q^{-m}.$$

Summing over all possible choices of $\{\epsilon_i\}_{i=1}^q$ such that $\sum_{i=1}^q \epsilon_i = s$, the proof is concluded.

3. Proof of Theorem 2

Let

$$M_k(q,d) := \frac{1}{|\mathcal{F}_d|}\sum_{F\in\mathcal{F}_d}\left(\frac{S(F)}{\sqrt{q}}\right)^k$$

be the kth moment of $S(F)$ as F ranges over the family of square-free polynomials of degree d in $\mathbb{F}_q[X]$. As before, let $X_1,\dots,X_q$ be independent trinomial random variables, taking values -1, 0, 1 with probabilities $(\frac{q/2}{q+1}, \frac{1}{q+1}, \frac{q/2}{q+1})$. Theorem 2 is then an immediate consequence of the following proposition.

Proposition 7. *We have*

$$M_k(q,d) = \mathbb{E}\left(\left(\frac{1}{q^{1/2}}\sum_{i=1}^{q} X_i\right)^k\right) + O\left(q^{(3k-d)/2}\right).$$

In particular, if $q,d\to\infty$, $M_k(q,d)$ agrees with Gaussian moments for all k.

Proof. We have

$$\begin{aligned} M_k(q,d) &= \frac{1}{|\mathcal{F}_d|}\sum_{f\in\mathcal{F}_d}\left(\frac{1}{q^{1/2}}\sum_{x\in\mathbb{F}_q}\chi\big(f(x)\big)\right)^k = \frac{1}{q^{k/2}}\sum_{x_1,x_2,\dots,x_k\in\mathbb{F}_q}\ \sum_{f\in\mathcal{F}_d}\chi\big(f(x_1)f(x_2)\cdots f(x_k)\big)\\ &= \frac{1}{q^{k/2}}\sum_{l=1}^{k} c(k,l)\sum_{((x_1,\dots,x_l),(\epsilon_1,\dots,\epsilon_l))\in P_{k,l}}\frac{1}{|\mathcal{F}_d|}\sum_{f\in\mathcal{F}_d}\chi\left(\prod_{i=1}^{l} f(x_i)^{\epsilon_i}\right) \end{aligned} \tag{3.1}$$

where

$$P_{k,l} = \left\{\big((x_1,\dots,x_l),(\epsilon_1,\dots,\epsilon_l)\big)\colon\ x_1,\dots,x_l \text{ all distinct and } \sum_{i=1}^{l}\epsilon_i = k\right\}$$

and $c(k,l)$ is a certain combinatorial factor, whose exact form is unimportant. Now, by Lemma 5,

$$\frac{1}{|\mathcal{F}_d|}\sum_{f\in\mathcal{F}_d}\chi\left(\prod_{i=1}^{l} f(x_i)^{\epsilon_i}\right) = 0 + O\left(q^{l-d/2}\right)$$

unless *all* ϵ_i are even, in which case

$$\frac{1}{|\mathcal{F}_d|}\sum_{f\in\mathcal{F}_d}\chi\left(\prod_{i=1}^{l} f(x_i)^{\epsilon_i}\right) = \frac{1}{(1+1/q)^l} + O\left(q^{l-d/2}\right).$$

Noting that

$$\sum_{l=1}^{k} c(k,l)\sum_{((x_1,\dots,x_l),(\epsilon_1,\dots,\epsilon_l))\in P_{k,l}} 1 = q^k$$

we find that the contribution from the error terms is $\ll q^{-k/2}q^k q^{k-d/2}$ and thus (3.1) equals

$$\frac{1}{q^{k/2}}\sum_{l=1}^{k} c(k,l) \sum_{\substack{((x_1,\dots,x_l),(\epsilon_1,\dots,\epsilon_l))\in P_{k,l}\\ \text{all } \epsilon_i \text{ even}}} \frac{1}{(1+1/q)^l} + O\left(q^{(3k-d)/2}\right).$$

On the other hand, since $X_1,\dots,X_q$ are independent trinomial random variables, we find that the expectation of the kth moment of their normalized sum is

$$\begin{aligned}\mathbb{E}\left(\left(\frac{1}{q^{1/2}}\sum_{j=1}^{q} X_j\right)^k\right) &= \frac{1}{q^{k/2}} \sum_{i_1,i_2,\dots,i_k\in\{0,1,\dots,q-1\}} \mathbb{E}(X_{i_1}\cdot X_{i_2}\cdots X_{i_k})\\ &= \frac{1}{q^{k/2}}\sum_{l=1}^{k} c(k,l) \sum_{((x_1,\dots,x_l),(\epsilon_1,\dots,\epsilon_l))\in P_{k,l}} \mathbb{E}\left(\prod_{j=1}^{l} X_{i_j}^{\epsilon_i}\right).\end{aligned}$$

As before we have $\mathbb{E}(\prod_{j=1}^{l} X_{i_j}^{\epsilon_i}) = 0$, unless all ϵ_i are even, in which case

$$\mathbb{E}\left(\prod_{j=1}^{l} X_{i_j}^{\epsilon_i}\right) = \frac{1}{(1+1/q)^l}$$

(note that $\mathbb{E}(X_{i_j}^{\epsilon}) = 0$ for ϵ odd, and $\mathbb{E}(X_{i_j}^{\epsilon}) = 1/(1+q^{-1})$ for ϵ even), concluding the proof of the first assertion.

The final assertion follows since the moments of a sum of bounded i.i.d. random variables converge to the Gaussian moments, cf. [2, Section 30]. □

References

[1] Jonas Bergström, Equivariant counts of points of the moduli spaces of pointed hyperelliptic curves, preprint, http://arxiv.org/abs/math/0611813v1, 2006.

[2] Patrick Billingsley, Probability and Measure, third ed., Wiley Ser. Probab. Math. Stat., John Wiley & Sons Inc., New York, 1995, A Wiley–Interscience Publication.

[3] Persi Diaconis, Mehrdad Shahshahani, On the eigenvalues of random matrices, in: Studies in Applied Probability, J. Appl. Probab. 31A (1994) 49–62.

[4] Dmitry Faifman, Zeev Rudnick, Statistics of the zeros of zeta functions in families of hyperelliptic curves over a finite field, preprint, http://arxiv.org/abs/0803.3534, 2008.

[5] Nicholas M. Katz, Peter Sarnak, Random Matrices, Frobenius Eigenvalues, and Monodromy, Amer. Math. Soc. Colloq. Publ., vol. 45, American Mathematical Society, Providence, RI, 1999.

[6] L.A. Knizhnerman, V.Z. Sokolinskii, Some estimates for rational trigonometric sums and sums of Legendre symbols, Uspekhi Mat. Nauk 34 (3 (207)) (1979) 199–200.

[7] L.A. Knizhnerman, V.Z. Sokolinskii, Trigonometric sums and sums of Legendre symbols with large and small absolute values, in: Investigations in Number Theory, Saratov. Gos. Univ., Saratov, 1987, pp. 76–89.

[8] Michael Larsen, The normal distribution as a limit of generalized Sato–Tate measures, preprint, arXiv:0810.2012v1 [math.NT], 2008.